A Survey of Network Coverage Prediction Mechanisms in 4G Heterogeneous Wireless Networks

Fatema Shaikh
Sir Isaac Newton School of Technology
British Institute of Technology and Ecommerce
London, United Kingdom
fshaikh@bite.ac.uk

Glenford Mapp, Aboubaker Lasebae
School of Engineering and Information Sciences
Middlesex University
London, United Kingdom
(g.mapp, a.lasebae)@mdx.ac.uk

Abstract— Seamless connectivity in 4G wireless networks requires the development of intelligent proactive mechanisms for efficiently predicting vertical handovers. Random device mobility patterns further increase the complexity of the handover process. Geographical topologies such as indoor and outdoor environments also exert additional constraints on network coverage and device mobility. The ability of a device to acquire refined knowledge about surrounding network coverage can significantly affect the performance of vertical handover prediction and QoS management mechanisms. This paper presents a comprehensive survey of research work conducted in the area of 4G wireless network coverage prediction for the optimisation of vertical handovers. It discusses different coverage prediction approaches and analyses their ability to accurately predict network coverage.

Keywords-4G wireless networks, proactive vertical handovers.

I. INTRODUCTION

The vision of Fourth Generation (4G) heterogeneous networking is the provisioning of universal connectivity and mobility through the seamless integration of different network access technologies offering diverse levels of Quality of Service. Multi-interfaced mobile devices should roam freely among networks without experiencing disruptions like connection loss during handovers, while giving them the choice of the best available location-based network services. It is widely accepted that different networks in the heterogeneous environment will integrate in a loosely-coupled manner, with each network domain being independently deployed by a different service provider [1].

The vertical handover process consists of three well-defined phases – system discovery, handover decision and handover execution. Among these three phases, the handover decision phase is the most crucial and decisions taken in this phase can directly affect a Mobile Node's (MN) communication [2]. It aims to answer three fundamental questions about vertical handovers:

- *When?* – The quest for the answer to this question has resulted in the area of handover prediction.
- *Which?* – The answer to this question forms the area of network selection.
- *How much?* – The answer to this question is sought through resource allocation and QoS management techniques.

Correct decisions in this phase are mainly dependent on the refined tuning and blending of the correct answers to the above three questions. An effective way of minimizing/eliminating disruptions due to handovers is to equip the network and MN with the ability to proactively detect vertical handovers before they actually take place so that the devices can start procedures to prepare and adjust to impending changes in network conditions.

Correctness of decisions in the handover prediction phase to a great extent lies in the accuracy of the answer to the first question:

"When is the device expected to perform a vertical handover?"

This question requires refined knowledge of the extent of a network's availability and can largely affect the correctness and accuracy of decisions taken in response to the other two questions. An incorrect answer can lead to an overall degradation in performance due to instability in other phases of the vertical handover, even resulting in connection loss. An accurate knowledge of the duration of availability of a network in relation to a MN's motion within that network is crucial to the successful management of handover related issues in 4G heterogeneous networks.

II. AVOIDANCE OF UNNECESSARY VERTICAL HANDOVERS

In order to achieve seamless roaming in a wireless heterogeneous device, one important problem that needs to be eliminated is that of unnecessary vertical handovers. This means that the MN should remain connected to the new network for duration equal to the handover recovery period. This is the time in which the data received on the new interface is equivalent to at least the amount that would have been received on the old interface in the duration equal to the total handover procedure. Otherwise the handover will be considered as unnecessary if the MN is forced to perform a vertical handover once again before the recovery duration period expires.

The main causes of unnecessary handovers are the failure to recognise temporary coverage, unavailability of required resources and congestion in the new network. Among these, the problem of predicting temporary coverage still remains largely unresolved. For instance, A MN roaming into the strong but temporary coverage of a WLAN may have access to the most optimal resources and the most favourable channel conditions. Yet, the fact that it will have to perform an upward vertical handover before successfully utilising these resources means that their availability is virtually useless unless it is harnessed in the correct manner. An unnecessary vertical handover actually results in an increased signalling overhead and delay.

The possibility of a handover being unnecessary is dependent mainly on whether the new network can satisfy the requesting traffic stream's resource demands in the limited period of connectivity. For instance, a downward vertical handover will not be considered futile if the enqueued data in the MN is a set of emails which can easily be sent in the limited period of time that a MN connects to the temporary but free hotspot. However, the new connection may not be suitable for starting a VOIP connection and the handover in this case will be considered unnecessary.

III. IMPACT OF GEOGRAPHICAL TOPOLOGY

In the prediction of vertical handovers, a crucial piece of information that has so far been ignored by studies is the effect of geographical topologies and physical boundaries on the accuracy of handover prediction. This knowledge can affect the validity of a MN's decision to perform a vertical handover. Take the example of the scenario shown in figure 1 in which a MN is located inside a closed indoor environment and which moves towards the boundary of WLAN coverage in trajectory 3. According to a pure Received Signal Strength (RSS) based

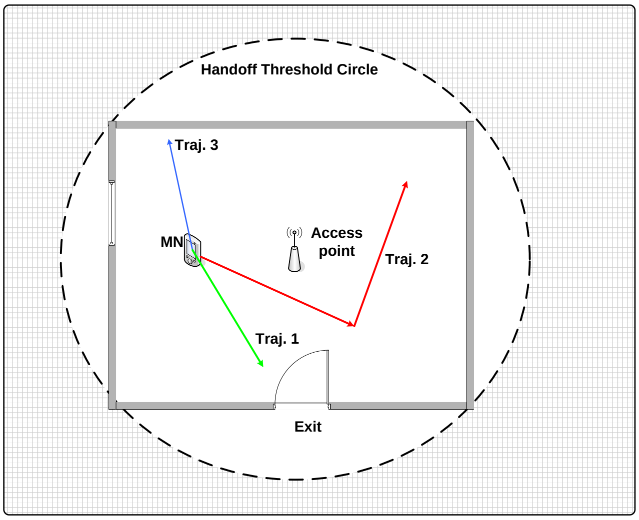

Figure 1. False handover triggers due to topological boundaries

handover prediction approach, rapidly decreasing RSS indicates that this MN is moving towards the coverage boundary, prompting the device to begin preparation for vertical handover. However if the network's coverage boundary threshold falls close to but beyond the environment's physical boundary like a wall, the reality is that the device cannot experience a vertical handover as it will be prevented from exiting the current coverage by the physical boundary. Scenarios like these are becoming increasingly common due to the widespread deployment of WLAN hotspots. Therefore a key requirement for seamless handovers is more detailed and refined knowledge of the geographical topology surrounding the MN.

IV. CLASSIFICATION OF COVERAGE PREDICTION MECHANISMS

In this section we classify a representative set of coverage prediction mechanisms into three popular categories:
1. History-based coverage prediction
2. Database-based coverage prediction
3. Mathematical modelling-based coverage prediction.

We then critically review each of these methods in the light of their suitability to meet the dynamic demands of intelligent coverage prediction in wireless heterogeneous clients.

A. History based coverage prediction

In literature, a number of studies adopted the history-based approach for improved coverage and context awareness in handover prediction mechanisms. The main assumption of this approach is that a MN's movement patterns in the future are mostly likely to be similar to past patterns and thus can be predicted from stored sequences of data.

Liu and colleagues [3] proposed a set of mobile motion prediction algorithms designed to predict the future location of a mobile user according to the user's movement history. The assumption was that users had a degree of regularity in their movements which could be recorded to predict their future movement patterns. The cellular system was divided into service areas, each area forming a set being represented by a state variable. MN movements were modelled by a discrete-parameter and discrete-state stochastic process consisting of states. This study recognised the importance of detecting coverage boundaries for improved context awareness. In an extension to this study, Stathes and Merakos [4] applied the concept of states to propose a path prediction algorithm based on learning automaton. The shortcoming of both these approaches was that they were limited to predicting the state of a MN within a service area and did not capture or predict its actual trajectory or refined movements within that service area.

Navidi and Liang in [5] and [6] proposed predictive distance-based mobility management schemes that attempted to predict the future location of a MN. While Liang based this on the Probability Density Function (PDF) of MN location given by a

Gauss-Markov model, Navidi did not assume any specific mobility model and suggested a history-based approach for location prediction based on previous reported locations. Both these approaches utilised context information mainly for reducing the paging overhead and did not delve deeply in the issue of refined coverage prediction.

Another study [7] proposed improved handover prediction for high priority users like rescue teams in very densely populated areas such as stadiums. The mechanism relied on large amounts of stored historical sequences for the development of a variety of mobility models which aimed to realistically model the behaviour of the MN in different situations like vehicular, pedestrian and group-dependent scenarios. Mobility prediction was limited to high priority users only and all available information was simply used to predict the next cell to which the MN was expected to perform a handover. Realising the importance of knowing the time until the next handover for smooth mobility management, the authors proposed calculating it as the exponentially smoothed mean based on all stored MN residence times. The problem with this technique was that even without considering random MN movement patterns, mean residence times could vary largely based on MN speeds.

Another study [8] applied the Markovian modelling approach to improve the predictability of the random walk model with the user moving between different states. Cell dwell time was history based and the approach did not capture the anomalies in handover prediction arising due to topological factors.

The main drawback of the history-based approach was that it relied heavily on large amounts of stored historical sequences. This was often insufficient to fully capture the random and spontaneous behaviour of MNs, e.g. pedestrian behaviour. Path prediction failed as soon as the MN strayed away from the predetermined route.

B. Coverage based handover prediction

Schemes from this handover prediction category employed previously stored knowledge about network coverage to predict the duration and quality of future coverage for a MN.

Soh et al. [9] proposed a proactive mobility prediction technique that applied both MN positioning information and road topology knowledge to predict the time a MN had before performing a horizontal handover. The study demonstrated how the knowledge of time before handover (TBH) helped in improving resource reservation efficiency and network performance. However, as this approach relied mainly on large volumes of data on road maps stored in prediction databases inside every BS, it was unable to predict the path and TBH of a MN when it strayed away from the road topology stored in the database. Thus the accuracy of TBH prediction decreased considerably when the MN exhibited random mobility patterns. Predicting the important role location and context/situation knowledge would play in the efficient integration of different technologies E. Cianca and colleagues [10] proposed a network-based middleware solution were intelligent agents in the access network provided support to users in tasks like network selection, resource and handover management, service discovery and QoS parameter adaptation. Being network controlled however, this approach was not active enough to react to sudden changes in network conditions at MN interfaces.

A recent study conducted in the area of handover management [11] employed coverage maps for improved network coverage prediction exclusively in vehicular environments. A noteworthy achievement of this study was the presence of real-time coverage data gathered by diligently driving a vehicle around a city and measuring the received signal strengths of different detected networks. However the study itself acknowledged the lengthy duration of the data gathering phase as a year and half long. Long collections phases may not always be feasible and database records can get outdated after a few years.

The drawback of the coverage database approach was that it involved the storage of large amounts of coverage information which automatically introduced increased overhead associated with the data gathering phase during the construction of coverage databases. In order to remain functional the databases required the frequent updating of coverage data which by no means is a simple task when considering large coverage areas like cities.

C. Problems with RSS as sole trigger for handovers

Many studies in literature adopted Received Signal Strength (RSS) as a key indicator of network availability. In wireless networks, although a rapidly deteriorating value of RSS can be a good indicator that the MN is approaching the coverage boundary and may soon perform an imminent handover (horizontal or vertical), in heterogeneous networking the metric alone cannot be considered a reliable trigger due to the following reasons:

- The RSS from different networks varies significantly because of differences in coverage and differences in techniques employed at the physical layers due to which they cannot be easily compared [12]. RSS fading patterns can also be very different due to large differences in BS-MN distances for different networks.

- RSS measurements alone cannot provide answers to complex questions such as the precise knowledge of how long the MN is expected to remain in the access point’s (AP) coverage. This knowledge is important for decision-making during both horizontal and vertical handovers as it can at a very early stage allow the MN to take important decisions on matters of resource allocation and QoS management.

- Rapid variations in signal levels due to phenomena such as multipath fading and shadowing and sudden changes in MN speeds and directions make it difficult to predict future RSS and signal quality [12].

- A MN employing RSS as a handover trigger is programmed to scan available channels once the RSS from the current AP falls below a threshold. However if a MN is powered up at the border of a set of WLAN cells with measured RSS approaching the handover threshold, it will keep scanning for new APs [13] despite the fact that the APs currently available may be able to fulfill its resource demands, resulting in unnecessary vertical handovers.

Therefore what is needed is a more robust and proactive metric that not only gives the current status of network coverage availability but which can also predict for how long the coverage and network services are likely to remain available.

D. Mathematical and modelling based handover prediction

The mathematical modelling based handover prediction category consists of theories that aim to predict future handover conditions by dynamically applying mathematically derived formulae and models to available network information.

Ylianttila et al [14] developed an extension to the dwell timer [16] scheme which was the predefined time for which a MN remained in the old network taking samples of the RSS from the AP and comparing them with a predefined threshold. If these samples taken in the dwell time were below the threshold then the MN initiated the handoff to the other network. Ylianttila;s study further proposed using a predefined dwell-timer for different data rates. While the study proposed intelligent solutions to eliminate the ping-pong effect, it did not provide a quantitative measure of the time the MN was expected to dwell in the old network. It also suffered from the inaccuracy resulting from employing only RSS as the threshold parameter. Bing H. et al in their study [16] demonstrated how when distance criterion was taken into account, the handover probabilities are smaller than those only based on RSS criterion.

H. Wang and colleagues [17] defined the useful concept of stability period which was the minimum duration for which the new network had to consistently display the lowest value of fn in order to make it better choice for handover. This period was

$$Tmakeup + Thandoff$$

$Tmakeup$ was the amount of time required to make up for loss of data or money due to handover latency and $Thandoff$ was the handover latency itself. Values of these two parameters were based on recent measurements in the past. The study also assumed the availability of context information such as network maps which could aid in the deduction of their values.

Chen et al. [18] further refined the stability period through the utility function which was the sum of the product of network parameters and their assigned weights. The utility ratio was the ratio of target network utility by current network utility. The duration of stability period increased or decreased dynamically based on the decrease or increase in utility ratio. Both Chen and Wang focused on deriving a quantitative measure of the time needed to overcome the effect of a vertical handover which in turn would decide whether it was suitable to switch to the new network. However the studies did not predict how long the MN was expected to reside in the current network coverage.

Zhu and McNair [19] proposed a policy-based vertical handoff decision algorithm where the calculated cost provided a measurement of the benefit of handing off to a certain network. However elimination parameters considered were RSS and channel availability only which do not provide sufficient information if the MN wants to avoid unnecessary vertical handovers.

In order to solve incertitude in vertical handovers due to the Line of Sight vulnerability in 60GHz LOS interfaces, Wang and colleagues [20] proposed an algorithm based on Decision Theory which aimed to predict the duration of disruption in LOS communication which in turn helped to decide whether or not the device should switch to WLAN. The scope of this approach was limited to resolving incertitude in LOS communication in indoor environments and it did not consider issues arising in WLAN based vertical handovers due to topological effects.

X. Yan et al [21] proposed a mathematically derived model based on the prediction of travelling distance which aimed to avoid unnecessary vertical handovers from cellular networks to WLANs. The proposed approach was RSS-based for MN-AP distance calculation and was based on the assumption of a circular WLAN coverage. While this approach did succeed in predicting unnecessary vertical handovers to WLAN, it had several shortcomings. First the solution could predict an unnecessary vertical handover only when the MN's trajectory actually cut the WLAN cell coverage boundary and could not capture the random movements of the MN within the WLAN cell. For example the technique would not work if a MN entered the WLAN cell, stopped, changed direction and moved inwards towards the centre. Secondly, the accuracy of the proposed solution in predicting unnecessary vertical handovers increased only for MNs travelling with speeds above 15 m/s which is rather unrealistic as previous results have demonstrated that the accuracy of the system decreased by up to 70% when MN speeds fell below 10 m/s. A reason for high accuracy in higher speeds was that the faster the MN travelled, the lesser was the RSS sampling rate.

On analysing the performance of the mathematical modelling approach, the key challenge that emerges is that the process is computationally intensive so it is important to develop efficient solutions which do not exert great demands on the MN's limited computational resources.

V. SURVEY OUTCOMES AND CONCLUSION

One of the main outcomes of the survey conducted in this paper is the revelation that despite the availability of a rich variety of context information, wireless heterogeneous devices still lack the intelligence to recognise their surrounding environment, particularly the precise knowledge of network coverage availability. The review evaluated a number of studies that proposed various solutions to tackle the issue of predicting network coverage. However none of them succeeded in providing a simple and effective solution that dynamically predicts the duration of network coverage availability for a MN. In other words they fail to dynamically provide a flexible answer to the question:

How can a MN roaming within a network predict future network availability relative to its motion within the network, and determine how long it has before it performs a vertical handover?

An important reason for this failure is the lack of new type of context information that specifically recognises coverage boundaries. While it may be argued that coverage based prediction techniques do provide a means to obtain this information, the approach does not facilitate the dynamic detection of boundaries by a MN which is a crucial requirement of 4G clients and requires the continuous supply of coverage data. The second important deficiency that emerged common in almost all approaches was their inability to accommodate the truly random movement behaviour of MNs, particularly in pedestrian environments. The third deficiency unravelled was the failure to empower the MN with proactive mechanisms that enabled it to calculate network coverage availability dynamically as per its needs. The fourth deficiency was the lack of a hybrid mechanism that functioned correctly while considering both vehicular and pedestrian speeds and behaviour. This was largely because most studies chose RSS as the key decision parameter for handovers. All these deficiencies form the key gaps in knowledge that need to be addressed in order to improve the correctness and accuracy of decisions made during the handover prediction phase.

[1] Song Q and Jamalipour, A. Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques. *IEEE Wireless Communications* Vol. 12, 42-48. 2005.

[2] Persson, H. and Karlsson, J.M. An Approach to Structure the Handover Terminology. *Proceedings of SNCNW'05.* 2005.

[3] Liu, G. and Maguire, G.J. A Class of mobile motion prediction algorithms for wireless mobile computing and communications. *Mobile Networks and Applications* Vol. 1, 113-121, 1996.

[4] Hadjiefthymiades, S. and Merakos, L. Proxies + Path Prediction: Improving Web Service Provision in Wireless-Mobile Communications. *Mobile Networks and Applications,* Vol. 8, 389-399. Kluwer Academic Publishers. 2003.

[5] Navidi, W. and Camp, T. Predicting node location in a PCS network. *Proceedings of the International Conference on Performance, Computing and Communications* 165-170. 2004.

[6] Liang, B. and Haas, Z.J. Predictive distance-based mobility management for multi-dimensional PCS networks. *IEEE/ACM Transactions on Networking* Vol. 11, Issue 5, 718-732, October 2003.

[7] Michaelis, S., Lewandowski, A., Daniel, K., Yousaf, F.Z. and Wietfeld, C. A Comprehensive Mobility Management Solution for Handling Peak Load in Cellular Network Scenarios. *Proceedings of the 6th ACM International Symposium on Mobility Management and Wireless Access* , 9-18. 2008.

[8] Fulop, P., Szalka, T., Szabo, S., Lendvai, K. and Imre, S. Accurate mobility modelling and location prediction based on pattern analysis of handover series in mobile networks. *Proceedings of MoMM2008,* 219-226. 24-26 November 2008.

[9] Soh, W. and Kim, H., S. Dynamic Bandwidth Reservation in Cellular Networks using Road Topology Based Mobility Predictions. *Proceedings of the 23rd IEEE INFOCOM Conference,* Vol. 4, 2766-2777. March, 2004.

[10] Cianca, E., Molinaro, A., Sanctis, M.D., Iera, A., Araniti, G. and Ruggieri, M. Location/Situation-Aware Architecture for Mobility Management over Heterogeneous Networks. *Proceedings of the 2nd international conference on Mobile multimedia communications.* 2006.

[11] Cottingham, D. Vehicular Wireless Communication. UCAM-CL-TR-741, January 2009.

[12] Makela, J., Ylianttila, M. and Pahlavan, K. Handoff decision in multi-service networks. *The 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*. Vol. 1, 655-659, 2000.

[13] Liao, Y. and Gao, L. Practical Schemes for Smooth MAC Layer Handoff in 802.11 Wireless Networks. *Proceedings of the 2006 International Symposium on World of Wireless, Mobile and Multimedia Networks*, 181-190. 2006.

[14] Ylianttila, M., Pande, M., Makela, J. and Mahonen, P. Optimization scheme for mobile users performing vertical handoffs between IEEE 802.11 and GPRS/EDGE networks. *Proceedings of IEEE GLOBECOM'01,* Vol. 6, 3439-3443, 2001.

[15] Hatimi, A., Krishnamurthy, P., Pahlavan, K., Ylianttila, M., Makela, J. and Pichna, R. 1999. Analytical Framework for Handoff in Non-Homogeneous Mobile Data Networks. *Proceedings of PIMRC'99,* September 1999.

[16] Bing, H., Lingge, J., Performance Analysis of Vertical Handover in a UMTS-WLAN Integrated Network. *Proceedings of the 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication*. 2003.

[17] Wang, H.J., Katz, R.H. and Giese, J. Policy-enabled handoffs across heterogeneous wireless networks. *Proceedings of Second IEEE Workshop on Mobile Computing Systems and Applications,* 51-60.1999.

[18] Chen, W., Liu, J. and Huang, H. An Adaptive Scheme for Vertical Handoff in Wireless Overlay Networks. *Proceedings of the 10th International Conference on Parallel and Distributed Systems.* 2004.

[19] Mcnair, J. and Fang Zhu. Vertical handoffs in fourth-generation multi-network environments. *IEEE Wireless Communications* Vol. 11, Issue 3, 8-15. 2004.

[20] Wang, J., Prasad, R.V. and Niemegeers, I.G.M.M. Solving Incertitude of Vertical Handovers in Heterogenoeus Mobile Wireless Network. *Proceedings of WICON'07,* October 22-24, 2007.

[21] Xiaohuan Yan, Mani, N. and Cekercioglu, Y.A. A Traveling Distance Prediction Based Method to Minimize Unnecessary Handovers from Cellular Networks to WLANs. *IEEE Communications Letters* Vol. 12, 14-16. 2008.